

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/538,481  
Source: IFWO  
Date Processed by STIC: 8/23/06

# ***ENTERED***



IFWO

## RAW SEQUENCE LISTING

DATE: 08/23/2006

PATENT APPLICATION: US/10/538,481

TIME: 08:41:37

Input Set : A:\271245.sub.seq.list.081706.txt

Output Set: N:\CRF4\08232006\J538481.raw

3 <110> APPLICANT: Mitsukan Group Corporation  
 4 Goto, Hidetsugu  
 6 <120> TITLE OF INVENTION: Structural gene responsible for high temperature tolerance  
 in acetic  
 7 acid bacteria, acetic acid bacteria transformed with said gene, and  
 8 acetic acid fermentation using said transformants.  
 10 <130> FILE REFERENCE: 271245US0PCT  
 12 <140> CURRENT APPLICATION NUMBER: 10/538,481  
 C--> 14 <141> CURRENT FILING DATE: 2005-06-07  
 16 <150> PRIOR APPLICATION NUMBER: PCT/JP03/15542  
 18 <151> PRIOR FILING DATE: 2003-12-04  
 20 <150> PRIOR APPLICATION NUMBER: JAPAN 2002-356844  
 22 <151> PRIOR FILING DATE: 2002-12-09  
 24 <160> NUMBER OF SEQ ID NOS: 7  
 26 <210> SEQ ID NO: 1  
 27 <211> LENGTH: 1313  
 28 <212> TYPE: DNA  
 29 <213> ORGANISM: Gluconacetobacter entanii  
 31 <400> SEQUENCE: 1  
 33 gaagagtgat attacacttc cctgacgccg ttttctaatt tgctccatac gcggggacctt 60  
 34 gccggaaaga taatgtctgt tttcaacgct cttgtttcac ccgccggact ggccgcgacc 120  
 35 gttgcggtcg ccgggtgcat gcaggcagcg cttggcactt ttctggtctc gcgtttccgg 180  
 36 tggcaggaaa aacgcattgga ccgggcgggtg cccatgcctc cggtttccgt gctcaagccc 240  
 37 ctccacggcg atgaaccgct gctggaggaa gcgcttgaaa gcttctgcac gcaggattac 300  
 38 ccgcagatgc agatcgtctt tggcgtacag gccgaagacg atgcggcgat cccgatcgta 360  
 39 caacggttga tggaaaccca cccgatgtg cagatggaac tggtgattga cccaccttc 420  
 40 cacgggctca accgcaagat cggcaacctg atcaacatca tgacgcgcgt gaagcatgat 480  
 41 gtctctggtca tttccgattc ggatatccac gttgcccccg attacctgcg gcatgtggtg 540  
 42 ggcgccatgg tgcccgacaa tgcggcctg gtcacgacgc tgtacgcggg gctgcccgcg 600  
 43 tcatccacgc tgccgcgcct gctggccgca tgccagatca accataactt cctgcccggc 660  
 44 gtgatgctgt cactctacct cgggcggcag gactgccttg gggcgacaat ggcgctgcgg 720  
 45 cgttccatgc tggacgaaat cggcgggctg gaagccctcg tgccgcatgt ggccgatgat 780  
 46 gcgatactgg gccgttacgt gcgtgaccgt ggcaaggata tcgccattgc cgcgtgcatg 840  
 47 acctggacca ccgtgggcga gacctcgatg cgtgaggtgc tggcgcatga actgcgctgg 900  
 48 ggccggaccg tcaagacgct ggagcctgcg gggtatgccg catccgccat ccagctgcc 960  
 49 ctgttctggtg ccagcgtcgc cgtgcttgcc gcgccgcatg cgacctggac atggtccttc 1020  
 50 tttcttggtg catggggatg gcgggcccgtg tgttccttca tcctggaccg tacgctggcg 1080  
 51 caacgtagtc tgggtgctgcc gtactgctt ctgccactgc gcgactggat ctcgcccgcc 1140  
 52 gtcatggttg gcagtgtcac tggcacgcgg gttgcatggc gtgggcagac aatgcatgtc 1200  
 53 acgccccatt cggtcatgac accacgatcg caaccggctt cccccggtga ctgaccgcgc 1260  
 54 gtcagcagcg tgaactgctt gagaattcca accctgtcgt taataagaac ggg 1313  
 59 <210> SEQ ID NO: 2  
 60 <211> LENGTH: 393  
 61 <212> TYPE: PRT

## RAW SEQUENCE LISTING

DATE: 08/23/2006

PATENT APPLICATION: US/10/538,481

TIME: 08:41:37

Input Set : A:\271245.sub.seq.list.081706.txt

Output Set: N:\CRF4\08232006\J538481.raw

```

62 <213> ORGANISM: Gluconacetobacter entanii
64 <400> SEQUENCE: 2
66 Met Ser Val Phe Asn Ala Leu Val Ser Pro Ala Gly Leu Ala Ala Thr
67              5              10              15
68 Val Ala Val Ala Gly Cys Met Gln Ala Ala Leu Gly Thr Phe Leu Val
69              20              25              30
70 Ser Arg Phe Arg Trp Gln Glu Lys Arg Met Asp Arg Ala Val Pro Met
71              35              40              45
72 Pro Pro Val Ser Val Leu Lys Pro Leu His Gly Asp Glu Pro Leu Leu
73              50              55              60
74 Glu Glu Ala Leu Glu Ser Phe Cys Thr Gln Asp Tyr Pro Gln Met Gln
75 65              70              75              80
76 Ile Val Phe Gly Val Gln Ala Glu Asp Asp Ala Ala Ile Pro Ile Val
77              85              90              95
78 Gln Arg Leu Met Glu Arg His Pro Asp Val Gln Met Glu Leu Val Ile
79              100             105             110
80 Asp Pro Thr Phe His Gly Leu Asn Arg Lys Ile Gly Asn Leu Ile Asn
81              115             120             125
82 Ile Met Thr Arg Val Lys His Asp Val Leu Val Ile Ser Asp Ser Asp
83              130             135             140
84 Ile His Val Ala Pro Asp Tyr Leu Arg His Val Val Gly Ala Met Val
85 145             150             155             160
86 Pro Asp Asn Val Gly Leu Val Thr Thr Leu Tyr Ala Gly Leu Pro Ala
87              165             170             175
88 Ser Ser Thr Leu Pro Arg Leu Leu Ala Ala Cys Gln Ile Asn His Asn
89              180             185             190
90 Phe Leu Pro Gly Val Met Leu Ser Leu Tyr Leu Gly Arg Gln Asp Cys
91              195             200             205
92 Leu Gly Ala Thr Met Ala Leu Arg Arg Ser Met Leu Asp Glu Ile Gly
93              210             215             220
94 Gly Leu Glu Ala Leu Val Pro His Val Ala Asp Asp Ala Ile Leu Gly
95 225             230             235             240
96 Arg Tyr Val Arg Asp Arg Gly Lys Asp Ile Ala Ile Ala Ala Cys Met
97              245             250             255
98 Thr Trp Thr Thr Val Gly Glu Thr Ser Met Arg Glu Val Leu Ala His
99              260             265             270
100 Glu Leu Arg Trp Gly Arg Thr Val Lys Thr Leu Glu Pro Ala Gly Tyr
101              275             280             285
102 Ala Ala Ser Ala Ile Gln Leu Pro Leu Phe Trp Ala Ser Val Ala Val
103              290             295             300
104 Leu Ala Ala Pro His Ala Thr Trp Thr Trp Ser Phe Phe Leu Gly Ala
105 305             310             315             320
106 Trp Gly Trp Arg Ala Val Cys Ser Phe Ile Leu Asp Arg Thr Leu Ala
107              325             330             335
108 Gln Arg Ser Leu Val Leu Pro Ser Leu Leu Leu Pro Leu Arg Asp Trp
109              340             345             350
110 Ile Ser Ala Ala Val Met Val Gly Ser Val Thr Gly Thr Arg Val Ala
111              355             360             365
112 Trp Arg Gly Gln Thr Met His Val Thr Pro His Ser Val Met Thr Pro

```

## RAW SEQUENCE LISTING

DATE: 08/23/2006

PATENT APPLICATION: US/10/538,481

TIME: 08:41:37

Input Set : A:\271245.sub.seq.list.081706.txt

Output Set: N:\CRF4\08232006\J538481.raw

```

113      370      375      380
114 Arg Ser Gln Pro Ala Ser Pro Gly Asp
115 385      390      393
118 <210> SEQ ID NO: 3
119 <211> LENGTH: 30
121 <212> TYPE: DNA
123 <213> ORGANISM: Artificial Sequence
125 <220> FEATURE:
127 <223> OTHER INFORMATION: Primer 1
129 <400> SEQUENCE: 3
131 gaagagtgat attacacttc cctgacgccg      30
134 <210> SEQ ID NO: 4
136 <211> LENGTH: 26
138 <212> TYPE: DNA
140 <213> ORGANISM: Artificial Sequence
142 <220> FEATURE:
145 <223> OTHER INFORMATION: Primer 2
148 <400> SEQUENCE: 4
150 cccgttctta ttaacgacag ggttgg      26
153 <210> SEQ ID NO: 5
155 <211> LENGTH: 5734
157 <212> TYPE: DNA
159 <213> ORGANISM: Gluconacetobacter entanii (Acetobacter altoacetigenes MH-24)
161 <400> SEQUENCE: 5
163 catggggcgt cacccccagc ggccagcttg gctacctgat ggacagggcg ggccttctgc      60
164 aagccctcgg ccaactgccat ctgccgggat atgaggccaa atacgaaccg aaggaaaagc      120
165 gcaccttctg ctaccccacc cagaacgcca gcggtcgggc tgtgcagcca tgatcgccaa      180
166 cccctccctc ttcttgagca attcgggaaga gcgatttccg ccgactgaac acgtcgaaaa      240
167 tggcagtttt ccaccgaaaa aaggaaagga ccataggaaa ggattaatat cttattttta      300
168 tctagggggt tgccgattcg cgattttcgc tgggaaaccg ccaaaaatgg cttgccatta      360
169 ggtcgcacca catgcgacca taaagtgcga cagtgtgcga cctattcggc ccatatacag      420
170 aggttcccca catgcggaat gtcacccgtc tcaagaccgc caaagaccgc ctccgcgagg      480
171 accaagccga cctggtgaag caagcccttc tgcccttcgc agaggacgat ggaccgatgc      540
172 gggatgcggt cggacggctc tacgtccaga tcaagaacct caccacccca gaccccgga      600
173 ccacggagcc gttcgtcatg atccgtcccg ccagaatcg cgcggtcacc ctctggctgc      660
174 tgaagaacag taagcggccc atgaaggccg tggacgtatg gacgctgctg ttcgaccacc      720
175 tgtttcccca taccggccag atcatgctga cccgtgagga aatcgcgga aaagtcggta      780
176 tccgggtcaa cgaagttaca gccgtcatga acgagctggt gagcttcggc gcgattttct      840
177 ccgagcgcga gaaggtggcc ggaatgcgcg ggccgggctt cgcccgctac tacatgaacc      900
178 ggcatgtggc cgaggtcggc agccgcgcca cgcaggaaga acttgcccta atcccacgcc      960
179 ccggcgccaa gctggcagtc gtgcagggtg gcaaggctta acccatgaag gtttcggaac      1020
180 tggacgtggt cgacagcgcc aaggcggcac aagaccggtt ggtgcgggaa gaactgctgc      1080
181 aagcagcgca ggcgagcggc cacggccccg ccctcgctca tgcccgttcc gtcatagcca      1140
182 aggcgcgggc cgggcaggac gccaaaggctt aacggccccg ccctctcccg cctcgatccc      1200
183 ggcgggcctg tagcatctcc tgatgctcct tggcggtttt ggcccgctgc tcggcccgct      1260
184 ctttctcggc cgctgcggct cttaggcgct cttcggccag ccgcacccgc tcgtccatct      1320
185 gacgtttccg atctgcctcg gcatccttgg cggctcctgc cttcagccct ttgctgaaag      1380
186 ccatccactt attggcgggt ttctcggtt tctgctgtat cggcgggggc agccgggtcaa      1440
187 atgcctgggc caccctctcg aagccctcac gcatggcggt gacggcctgc gccagtttag      1500

```

## RAW SEQUENCE LISTING

DATE: 08/23/2006

PATENT APPLICATION: US/10/538,481

TIME: 08:41:37

Input Set : A:\271245.sub.seq.list.081706.txt

Output Set: N:\CRF4\08232006\J538481.raw

188	ccagggcgaa	atctatcacc	tgggcccgt	gggcggttctc	ggcccggata	cgccgggtgt	1560
189	gggttgccggt	cggggtctgg	tggcccttcc	gttccagagc	caccacattc	ggcccatgt	1620
190	gccgctctgg	aacgcggtct	agcccttct	ccgcattgct	ccggtgatct	atccgggct	1680
191	cttgcccagc	cgctctagc	gcggcattgg	caaggccgc	ccatagctgc	cggatttct	1740
192	tcacctcgtc	ggcggccttc	cccagtccca	tgcctgccc	cttcttgctg	gacagttcga	1800
193	tgggtgattt	gtctccaaag	gacagcttgc	catcgccccc	ccgctccacc	gtgcgggtgg	1860
194	tgggtcatgat	gtgcgcgtga	tgattccggg	cgtcgccctc	gtcaccgcga	agatgcacgg	1920
195	ccacgtccac	ggccacccc	taccgctgga	ccaactcacg	cgcgaaactg	tccgccagtt	1980
196	cggcccgcgtg	ctcgtctggt	agttcatgag	ggagggccac	aaccattcc	ctcccggtgc	2040
197	gggcgctcctt	gcgtttctct	gatecgtccg	cgtcattcca	caattccgaa	cggtcagcgg	2100
198	tgccaccccc	cggaaatgaaa	attgccttat	gggcaacgct	attctgcctg	gggctgtatt	2160
199	tgtgttcgtg	ccgctcaacc	tcgttggtca	aatcctcgcc	agcacgatac	gcagccgcag	2220
200	ccacaacgga	acgccttgcg	ctccggctga	tcgggttctg	ttctgcgcga	tagattgcc	2280
201	cggatcgagc	gcctaccttt	tggagttaaa	cgggggggttc	aggggggcga	agccaccatg	2340
202	acgcaggact	tgcatttggt	caagtcgtaa	ctgcgcctt	aatacctgac	ggcatcaagg	2400
203	gatattgtggt	attcgtttga	aacggaacgg	ctccacgggtg	aggatgat	gagcgatatt	2460
204	gcgaaagaga	ttgagaacgc	caaaaggatc	atagctgaac	agaaaaagcg	catcaaagat	2520
205	gcccagaagg	aagcagctaa	agcggaatca	aagttgaggg	accgtcagaa	ctacatcttg	2580
206	ggcggcgcac	tggtaaaact	tgcggaaaca	gatgaacggg	ccgtccgcac	tattgaaaca	2640
207	cttttgaagc	tggtggatcg	tccatcagac	cggaaaggct	ttgagggtgt	ttcccgctctc	2700
208	ccatccctct	ccctgccac	gcagccagca	ccggacaccg	gccatgagtg	aggcactgga	2760
209	agaagatccg	tttgaactgt	tcaaaagggt	cgaaaaaagc	ctgtccacgg	ccaccgccag	2820
210	catggagcgg	ctggccgccc	aacaagatgc	caggtgcaag	accatttcag	acgccgccgg	2880
211	aaaagcctct	aaattggccg	aggaagccgg	tgacaccttc	acagcatcca	agaggcgtct	2940
212	gatgatctgg	acggccctct	gcgcggctct	gctggtctgt	ggcgggtggt	tggcgggtta	3000
213	ttggtgga	caccgtgacg	gttgggcctc	tggcacggcc	cacgacgtct	aagaaacat	3060
214	tattatcatg	acattaacct	ataaaaaatag	gcgtatcacg	aggcccttcc	gtctcgcgcg	3120
215	tttcggtgat	gacggtgaaa	acctctgaca	catgcagctc	ccggagacgg	tcacagcttg	3180
216	tctgtaagcg	gatgccggga	gcagacaagc	ccgtcagggc	gcgtcagcgg	gtgttggcgg	3240
217	gtgtcggggc	tggtcttaact	atgcggcatc	agagcagatt	gtactgagag	tgaccatat	3300
218	gcgggtgtgaa	ataccgcaca	gatgcgtaag	gagaaaaatac	cgcatacagg	gccattcgcc	3360
219	attcaggctg	gcgaactgtt	gggaaggcgg	atcggtgcgg	gcctcttcgc	tattacgcca	3420
220	gctggcgaaa	gggggatgtg	ctgcaaggcg	attaagttgg	gtaacgccag	ggttttccca	3480
221	gtcacgacgt	tgtaaaacga	cggccagtg	caagcttgca	tgcctgcagg	tcgactctag	3540
222	aggatccccg	ggtaccgagc	tcgaattcgt	aatcatggtc	atagctgttt	cctgtgtgaa	3600
223	attgttatcc	gctcacaatt	ccacacaaca	tacgagccgg	aagcataaag	tgtaaagcct	3660
224	ggggtgccta	atgagtgagc	taactcacat	taattgcgtt	gcgctcactg	cccgttttcc	3720
225	agtcgggaaa	cctgtcgtgc	cagctgcatt	aatgaatcgg	ccaacgcgcg	gggagaggcg	3780
226	gtttgcgtat	tgggcgctct	tccgcttct	cgtcactga	ctcgctgcgc	tcggctcgttc	3840
227	ggctgcggcg	agcggtatca	gctcactcaa	aggcggtaat	acggttatcc	acagaatcag	3900
228	gggataacgc	aggaaagaac	atgtgagcaa	aaggccagca	aaaggccagg	aaccgtaaaa	3960
229	aggccgcgtt	gctggcggtt	ttccataggc	tccgcccccc	tgacgagcat	cacaaaaatc	4020
230	gacgctcaag	tcagaggtgg	cgaaacccga	caggactata	aagataccag	gcgtttcccc	4080
231	ctggaagctc	cctcgtgcgc	tctcctgttc	cgacctgccc	gcttaccgga	tacctgtccg	4140
232	cctttctccc	ttcgggaagc	gtggcgcttt	ctcaatgctc	acgctgtagg	tatctcagtt	4200
233	cgggtgtaggt	cgttcgctcc	aagctgggct	gtgtgcacga	acccccggtt	cagcccgacc	4260
234	gctgcgcctt	atccggtaac	tatcgtcttg	agtcacaacc	ggtaaagacac	gacttatcgc	4320
235	cactggcagc	agccactggt	aacaggatta	gcagagcgag	gtatgtaggc	ggtgctacag	4380
236	agttcttgaa	gtggtggcct	aactacggct	acactagaag	gacagtattt	ggtatctgcg	4440

## RAW SEQUENCE LISTING

DATE: 08/23/2006

PATENT APPLICATION: US/10/538,481

TIME: 08:41:37

Input Set : A:\271245.sub.seq.list.081706.txt

Output Set: N:\CRF4\08232006\J538481.raw

```

237 ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa 4500
238 ccaccgctgg tagcgggtggt ttttttggtt gcaagcagca gattacgcgc agaaaaaaag 4560
239 gatctcaaga agatcctttg atcttttcta cggggtctga cgctcagtgg aacgaaaact 4620
240 cacgttaagg gattttgggtc atgagattat caaaaaggat cttcacctag atccttttaa 4680
241 attaaaaatg aagtttttaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt 4740
242 accaatgctt aatcagtgag gcacctatct cagcgatctg tctatttcgt tcatccatag 4800
243 ttgcctgact ccccgctcgtg tagataacta cgatacggga gggcttacca tctggcccca 4860
244 gtgctgcaat gataccgcga gacccacgct caccggctcc agatttatca gcaataaacc 4920
245 agccagccgg aagggccgag cgcagaagtg gtcttgcaac tttatccgcc tccatccagt 4980
246 ctattaattg ttgccgggaa gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg 5040
247 ttgttgccat tgctacaggc atcgtggtgt cagcgtcgtc gtttggtatg gcttcattca 5100
248 gctccggttc ccaacgatca aggcgagtta catgatcccc catgttgtgc aaaaaagcgg 5160
249 ttagctcctt cggctctccg atcgttgtca gaagtaagtt ggccgcagtg ttatcactca 5220
250 tggttatggc agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg 5280
251 tgactggtga gtactcaacc aagtcattct gagaatagt tatgcggcga ccgagttgct 5340
252 cttgcccggt gtcaatacgg gataataccg cgccacatag cagaacttta aaagtgtca 5400
253 tcattgaaa acgttcttgc gggcgaaaac tctcaaggat cttaccgctg ttgagatcca 5460
254 ttcgatgtaa cccactcgtg cacccaactg atcttcagca tcttttactt tcaccagcgt 5520
255 ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg 5580
256 gaaatgttga atactcatat tcttcctttt tcaatattat tgaagcattt atcagggtta 5640
257 ttgtctcatg agcggatata tatttgaatg tatttagaaa aataaacaaa taggggttcc 5700
258 ggcgcacattt ccccgaaaag tgccacctga cgtc 5734

```

261 &lt;210&gt; SEQ ID NO: 6

262 &lt;211&gt; LENGTH: 30

263 &lt;212&gt; TYPE: DNA

264 &lt;213&gt; ORGANISM: Artificial sequence

266 &lt;220&gt; FEATURE:

268 &lt;223&gt; OTHER INFORMATION: Primer A

270 &lt;400&gt; SEQUENCE: 6

272 cgctgacgtc gtgggccgtg ccagaggccc 30

275 &lt;210&gt; SEQ ID NO: 7

276 &lt;211&gt; LENGTH: 30

277 &lt;212&gt; TYPE: DNA

278 &lt;213&gt; ORGANISM: Artificial sequence

280 &lt;220&gt; FEATURE:

282 &lt;223&gt; OTHER INFORMATION: Primer B

284 &lt;400&gt; SEQUENCE: 7

286 ggccaagacg tctgcagcat ggggcgtcac 30

VERIFICATION SUMMARY

DATE: 08/23/2006

PATENT APPLICATION: US/10/538,481

TIME: 08:41:38

Input Set : A:\271245.sub.seq.list.081706.txt

Output Set: N:\CRF4\08232006\J538481.raw

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date